

Multivariate normal approximation in geometric probability

Mathew D. Penrose*

Department of Mathematical Sciences, University of Bath,
Claverton Down, Bath BA2 7AY, England.

Andrew R. Wade†

Department of Mathematics, University of Bristol,
University Walk, Bristol BS8 1TW, England.

April 2007

Abstract

Consider a measure $\mu_\lambda = \sum_x \xi_x \delta_x$ where the sum is over points x of a Poisson point process of intensity λ on a bounded region in d -space, and ξ_x is a functional determined by the Poisson points near to x , i.e. satisfying an exponential stabilization condition, along with a moments condition (examples include statistics for proximity graphs, germ-grain models and random sequential deposition models). A known general result says the μ_λ -measures (suitably scaled and centred) of disjoint sets in \mathbb{R}^d are asymptotically independent normals as $\lambda \rightarrow \infty$; here we give an $O(\lambda^{-1/(2d+\epsilon)})$ bound on the rate of convergence. We illustrate our result with an explicit multivariate central limit theorem for the nearest-neighbour graph on Poisson points on a finite collection of disjoint intervals.

Key words and phrases: Multivariate normal approximation; geometric probability; stabilization; central limit theorem; Stein's method; nearest-neighbour graph.

AMS 2000 Mathematics Subject Classification: 60D05, 60F05, 60G57.

*e-mail: m.d.penrose@bath.ac.uk

†e-mail: Andrew.Wade@bris.ac.uk

1 Introduction

There has been considerable recent interest in providing central limit theorems (CLTs) for certain functionals in geometric probability defined on spatial Poisson point processes. Such functionals include those associated with random spatial graphs such as the minimal-length spanning tree or the nearest-neighbour graph, as well as with germ-grain models and random sequential packing models. These functionals are random variables given by sums of contributions from points of a Poisson point process in \mathbb{R}^d .

A natural extension to random *measures* may be provided by keeping track of the location of each contribution in \mathbb{R}^d . In this way one can obtain a random field indexed by test functions on \mathbb{R}^d or by subsets of \mathbb{R}^d . For example, one can consider the measure induced by a Poisson process with a point mass at each Poisson point equal to the distance to its nearest-neighbour; then a typical multivariate statistic induced by this measure is the vector of total edge-lengths of the nearest-neighbour graph on Poisson points over a finite collection of disjoint subsets of \mathbb{R}^d .

Under certain conditions, it is known [4,10,11] that the measures, appropriately scaled and centred, of disjoint sets (or of test functions with disjoint supports) are asymptotically distributed as independent normals in the large-intensity limit. The object of the present paper is to give bounds on rate of convergence; these bounds are the main contribution of the present paper. We illustrate our result with an application to the nearest-neighbour situation mentioned above.

The unifying concept of *stabilization* on Poisson points has proved a useful notion of local dependence in the context of geometric probability. This says, roughly speaking, that the contribution from a Poisson point is unaffected by changes to the configuration of Poisson points beyond a certain (random) distance.

The methodology of stabilization has been fruitfully employed, in various guises, to produce univariate CLTs and laws of large numbers for random quantities in many problems in geometric probability; see e.g. [4,8–11,13–16]. The techniques used in this context include a martingale method (see for instance [8], and [13] where the method is presented for general stabilizing functionals in geometric probability), the method of moments [4], and Stein’s method [16], which we employ in the present paper.

The multivariate case, in which several collections of random variables are considered, has also received some attention. Applications in geometric probability include, for example, the joint normality of certain random spatial graph functionals defined over a finite collection of disjoint regions in \mathbb{R}^d . There are potential applications to multivariate statistics, including nonparametric multi-sample tests (see e.g. [17]).

In the present paper, we employ a form of *Stein’s method* (see [18]), which has the advantage that it can provide rates of convergence in the CLT. In this context, Stein’s method is a useful tool for establishing normal approximations and CLTs for sums of weakly dependent random variables. In this paper, the weak dependency structure is provided by the concept of stabilization on Poisson points.

In the univariate case, the method yields normal approximation of the sum of a single collection of random variables that are ‘mostly independent’, i.e. exhibiting a local dependency structure. This structure may be captured using dependency graphs. This method was first used in the context of geometric probability by Avram and Bertsimas in [2] (using the normal approximation error bounds of [3]) to provide CLTs for certain random combinatorial structures that are locally determined in some sense, including the

j -th nearest-neighbour graph, and the Delaunay and Voronoi graphs.

Using the sharper normal approximation bounds of [5], more general results for univariate normal approximation based on Stein's method for random point measures were given by Penrose and Yukich in [16]. That paper is the foundation for the present work, which is its multivariate analogue.

Multivariate CLTs for random measures in geometric probability have recently been proved via the method of moments [4] and also the martingale method [10]. In particular, [10] also covers lattice processes (such as percolation), and does not require 'exponential' stabilization, and so admits a larger class of measures. The advantage of the results in the present paper is that information on rates of convergence is provided.

Beyond the context of geometric probability, multivariate central limit theory has been well studied. Related results include multivariate central limit theorems for sums of independent random variables given in [6]. In [7, 17], multivariate normal approximation bounds are given for sums of (locally) dependent random variables, often chosen in somewhat special ways, including certain statistics defined on random graphs. The results in the present paper have the advantage of being more generally applicable in geometric probability.

2 Main result

The basic setting follows that of [16]. Let $d \in \mathbb{N}$. As in [16], we consider marked point processes in \mathbb{R}^d for the sake of generality. Let $(\mathcal{M}, \mathcal{F}_{\mathcal{M}}, \mathbb{P}_{\mathcal{M}})$ be a probability space (the mark space). Let $\xi(x, s; \mathcal{X})$ be a measurable $[0, \infty)$ -valued function defined for all triples $(x, s; \mathcal{X})$, where $x \in \mathbb{R}^d$, $s \in \mathcal{M}$ are such that $(x, s) \in \mathcal{X}$, where $\mathcal{X} \subset \mathbb{R}^d \times \mathcal{M}$ is finite. When $(x, s) \in (\mathbb{R}^d \times \mathcal{M}) \setminus \mathcal{X}$, we abbreviate notation and write $\xi(x, s; \mathcal{X})$ instead of $\xi(x, s; \mathcal{X} \cup \{(x, s)\})$.

Given $\mathcal{X} \subset \mathbb{R}^d \times \mathcal{M}$, $a > 0$ and $y \in \mathbb{R}^d$, set $y + a\mathcal{X} := \{(y + ax, s) : (x, s) \in \mathcal{X}\}$, i.e. translation and scaling act only on the 'spatial' part of \mathcal{X} . For all $\lambda > 0$ let

$$\xi_{\lambda}(x, s; \mathcal{X}) := \xi(x, s; x + \lambda^{1/d}(-x + \mathcal{X})).$$

Thus ξ_{λ} is a 'scaled-up' version of ξ , defined on a scaled-up version of the (marked) point set \mathcal{X} dilated around x . We say that ξ is *translation invariant* if $\xi(x + y, s; y + \mathcal{X}) = \xi(x, s; \mathcal{X})$ for all $y \in \mathbb{R}^d$, all $(x, s) \in \mathbb{R}^d \times \mathcal{M}$ and all finite $\mathcal{X} \subset \mathbb{R}^d \times \mathcal{M}$. When ξ is translation invariant, the functional ξ_{λ} simplifies to $\xi_{\lambda}(x, s; \mathcal{X}) = \xi(\lambda^{1/d}x, s; \lambda^{1/d}\mathcal{X})$.

For $q \in [1, \infty]$, let $\|\cdot\|_q$ denote the ℓ_q norm on \mathbb{R}^d . In the sequel we will use $q = 2$ (the Euclidean norm) and $q = \infty$.

Let κ be a probability density function on \mathbb{R}^d with compact support $A \subset \mathbb{R}^d$, where A is non-null (i.e. has non-zero Lebesgue measure). We assume throughout that κ is bounded with supremum denoted by $\|\kappa\|_{\infty} < \infty$. For all $\lambda > 0$ let \mathcal{P}_{λ} denote a Poisson point process in $\mathbb{R}^d \times \mathcal{M}$ with intensity measure $(\lambda\kappa(x)dx) \times \mathbb{P}_{\mathcal{M}}(ds)$.

We use the following notion of exponential stabilization, as given in [16] (taking the A_{λ} there to be A for all λ). For $x \in \mathbb{R}^d$ and $r > 0$, let $B_r(x)$ denote the Euclidean ball centred at x of radius r . Let U denote a random element of \mathcal{M} with distribution $\mathbb{P}_{\mathcal{M}}$, independent of \mathcal{P}_{λ} .

Definition 2.1 ξ is exponentially stabilizing with respect to κ and A if for all $\lambda \geq 1$ and all $x \in A$, there exists an almost surely finite random variable $R := R(x, \lambda)$, (a radius of stabilization for ξ at x) such that

$$\xi_\lambda(x, U; [\mathcal{P}_\lambda \cap (B_{\lambda^{-1/d}R}(x) \times \mathcal{M})] \cup \mathcal{X}) = \xi_\lambda(x, U; \mathcal{P}_\lambda \cap (B_{\lambda^{-1/d}R}(x) \times \mathcal{M})),$$

for all finite $\mathcal{X} \subset (A \setminus B_{\lambda^{-1/d}R}(x)) \times \mathcal{M}$, and moreover

$$\limsup_{t \rightarrow \infty} t^{-1} \log \left(\sup_{\lambda \geq 1, x \in A} \mathbb{P}[R(x, \lambda) > t] \right) < 0.$$

Roughly speaking, $R(x, \lambda)$ is a radius of stabilization if the value of ξ_λ at x is unaffected by changes to the configuration of Poisson points outside $B_{\lambda^{-1/d}R}(x)$. Exponential stabilization is known to hold for many ‘locally determined’ functionals defined on spatial point processes, and in particular in several cases of interest in geometric probability; see for example [16]. Following [16], we also make the following definition.

Definition 2.2 ξ has a moment of order $p > 0$ (with respect to κ and A) if

$$\sup_{\lambda \geq 1, x \in A} \mathbb{E} [|\xi_\lambda(x, U; \mathcal{P}_\lambda)|^p] < \infty. \quad (2.1)$$

For $\lambda > 0$, we define the random weighted point measure μ_λ^ξ on \mathbb{R}^d , induced by ξ_λ , by

$$\mu_\lambda^\xi := \sum_{(x,s) \in \mathcal{P}_\lambda \cap (A \times \mathcal{M})} \xi_\lambda(x, s; \mathcal{P}_\lambda) \delta_x,$$

where δ_x is the point measure at $x \in \mathbb{R}^d$.

For $\Gamma \subset \mathbb{R}^d$, let $\mathcal{B}(\Gamma)$ denote the set of bounded Borel-measurable functions on Γ . For $f \in \mathcal{B}(\Gamma)$, let $\langle f, \mu_\lambda^\xi \rangle := \int_\Gamma f d\mu_\lambda^\xi$. Let Φ denote, as usual, the standard normal distribution function on \mathbb{R} . We recall the following univariate normal approximation result of Penrose and Yukich (contained in Theorem 2.1 of [16]).

Proposition 2.1 [16] Let ξ be exponentially stabilizing and satisfy the moment condition (2.1) for some $p > 3$. For Γ a non-null Borel subset of A , let $f \in \mathcal{B}(\Gamma)$ and put $T := \langle f, \mu_\lambda^\xi \rangle$. Then there exists a constant $C \in (0, \infty)$ depending on d , ξ , f , and κ such that for all $\lambda \geq 2$,

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P} \left[\frac{T - \mathbb{E}[T]}{(\text{Var}[T])^{1/2}} \leq t \right] - \Phi(t) \right| \leq C(\log \lambda)^{3d} \lambda (\text{Var}[T])^{-3/2}. \quad (2.2)$$

For fixed $m \in \mathbb{N}$, let Γ_i , $i = 1, \dots, m$ be non-null Borel subsets of $A \subset \mathbb{R}^d$. For notational simplicity, for $i = 1, \dots, m$ and for $f_i \in \mathcal{B}(\Gamma_i)$ set $T_i := \langle f_i, \mu_\lambda^\xi \rangle = \int_{\Gamma_i} f_i d\mu_\lambda^\xi$. These are the quantities of interest to us in the present paper. By Proposition 2.1, under appropriate conditions, we have that, individually, each T_i satisfies a normal approximation result of the form of (2.2). For the present paper, we will impose one extra condition to control variances such as $\text{Var}[T_i]$.

(A1) There exist constants $C_i \in (0, \infty)$ such that for each i , for all λ sufficiently large, $\text{Var}[T_i] \geq C_i \lambda$.

Under assumption (A1), the bound on the rate of convergence on the right of (2.2) (in the case $T = T_i$) becomes $O(\lambda^{-1/2}(\log \lambda)^{3d})$ (compare Corollary 2.1 of [16]), and in particular (2.2) yields the central limit theorems

$$\frac{T_i - \mathbb{E}[T_i]}{(\text{Var}[T_i])^{1/2}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1),$$

as $\lambda \rightarrow \infty$, where $\mathcal{N}(0, 1)$ is the standard normal distribution on \mathbb{R} and ‘ $\xrightarrow{\mathcal{D}}$ ’ denotes convergence in distribution. Condition (A1) is true in many cases. In Section 4.1 we will give some sufficient conditions for (A1) to hold, and discuss alternative conditions which lead to somewhat stronger versions of (A1). In particular, it is often possible to show (under appropriate conditions) that $\lambda^{-1}\text{Var}[T_i] \rightarrow \sigma_i^2$ for some $\sigma_i^2 \in (0, \infty)$, which may be ‘explicit’ (see Section 4.1).

Our main result, Theorem 2.1 below, extends Proposition 2.1 to give a multivariate central limit theorem for $(T_i : i = 1, \dots, m)$, centred and scaled, with a bound on the rate of convergence. We impose the additional assumptions that (A1) holds and that the sub-regions Γ_i are pairwise disjoint and satisfy the natural regularity condition (A2) below. The central difficulty in extending Proposition 2.1 to a multivariate version is that the T_i are not, in general, independent. However, with the aid of stabilization we will show that they are ‘asymptotically independent’ in an appropriate sense.

To state (A2), we introduce some notation. For measurable $B \subset \mathbb{R}^d$, let $|B|$ denote the (d -dimensional) Lebesgue measure of B . Let ∂B denote the boundary of B . For $B \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$ let $d_q(x, B) := \inf_{y \in B} \|x - y\|_q$. Also, for $B, B' \subset \mathbb{R}^d$ with $B \cap B' = \emptyset$, let $d_q(B, B') := \inf_{x \in B, y \in B'} \|x - y\|_q$, i.e. the shortest distance (in the ℓ_q sense) between B and B' . For $r > 0$, let $\partial_r(B)$ denote the r -neighbourhood of the boundary of $B \subset \mathbb{R}^d$ in the ℓ_∞ norm, that is the set $\{x \in \mathbb{R}^d : d_\infty(x, \partial B) \leq r\}$.

(A2) For each i , $|\partial_r(\Gamma_i)| = O(r)$ as $r \downarrow 0$.

Sufficient conditions for (A2) include that each of the Γ_i is convex, or each is the finite union of convex regions (e.g. polyhedral). We can now state our main result.

Theorem 2.1 *Let ξ be exponentially stabilizing and satisfy the moment condition (2.1) for all $p \geq 1$. Let $m \in \mathbb{N}$. Let $\Gamma_1, \Gamma_2, \dots, \Gamma_m$ be fixed disjoint non-null Borel subsets of A satisfying (A2). For $i = 1, \dots, m$, let $f_i \in \mathcal{B}(\Gamma_i)$ and set $T_i := \langle f_i, \mu_\lambda^\xi \rangle$. Suppose that (A1) holds. Let $\varepsilon > 0$. Then there exists a constant $C \in (0, \infty)$ depending on $d, \xi, \kappa, \varepsilon, \{f_i\}$ and $\{\Gamma_i\}$, such that, for all $\lambda \geq 1$,*

$$\sup_{t_1, \dots, t_m \in \mathbb{R}} \left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ \frac{T_i - \mathbb{E}[T_i]}{(\text{Var}[T_i])^{1/2}} \leq t_i \right\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \leq C \lambda^{-1/(2d+\varepsilon)}. \quad (2.3)$$

In particular, from (2.3) we obtain the multivariate central limit theorem that says

$$\left(\frac{T_i - \mathbb{E}[T_i]}{(\text{Var}[T_i])^{1/2}} : i = 1, \dots, m \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, I_m), \quad (2.4)$$

as $\lambda \rightarrow \infty$, where $\mathcal{N}(0, I_m)$ is the m -dimensional normal distribution with mean 0 and covariance matrix given by the identity matrix I_m . It was already known [10, 11] that under

similar conditions to those of Theorem 2.1 we have (2.4), at least when $\lambda^{-1}\text{Var}[T_i] \rightarrow \sigma_i^2$ for some $\sigma_i^2 \in (0, \infty)$. Theorem 2.1 adds to this by providing a bound on the rate of convergence.

As an example of the application of Theorem 2.1, one can take $f_i = \mathbf{1}_{\Gamma_i}$ for $i = 1, 2, \dots, m$, where $\mathbf{1}_{\Gamma}$ is the indicator function of $\Gamma \subset \mathbb{R}^d$. We indicate some particular applications of Theorem 2.1 in Section 4. Under additional technical conditions, one can say more about the asymptotic behaviour of the variance terms in (2.3); see Section 4.1 below.

Remark. The relatively slow rate of convergence in higher dimensions arises primarily due to the possibility of strongly dependent points in the neighbourhood of the interface of adjacent regions. If all of the Γ_i are separated by a strictly positive distance, then our methods can be adapted to yield a rate of convergence of the same order as in the univariate result (Proposition 2.1), that is $O(\lambda^{-1/2}(\log \lambda)^{3d})$.

For ease of presentation, we prove Theorem 2.1 in Section 3 under the conditions that ξ is translation invariant and that the mark space is degenerate (i.e. $\mathcal{M} = \{1\}$), and so from now on we suppress any mention of \mathcal{M} . In particular, point sets such as \mathcal{X} and \mathcal{P}_λ will be treated as (their corresponding) subsets of \mathbb{R}^d , and we will write $\xi_\lambda(x; \mathcal{X})$ rather than $\xi_\lambda(x, 1; \mathcal{X})$. The proof can be adapted for the general marked case, as in [16].

3 Towards a proof of Theorem 2.1

For everything that follows, we assume that $\Gamma_1, \Gamma_2, \dots, \Gamma_m$ are (arbitrary) non-null Borel subsets of the bounded region $A \subset \mathbb{R}^d$, such that $\Gamma_i \cap \Gamma_j = \emptyset$ for $i \neq j$, and condition (A2) holds. Also, for each i we have a function $f_i \in \mathcal{B}(\Gamma_i)$.

For fixed $\alpha > 0$, let $s_\lambda := \alpha \lambda^{-1/d} \log \lambda$, and let Γ_i^{bd} denote the s_λ ‘boundary region’ of $\Gamma_i \subseteq A$, in the sense

$$\Gamma_i^{\text{bd}} := \{x \in \Gamma_i : d_\infty(x, \partial\Gamma_i) \leq \alpha \lambda^{-1/d} \log \lambda\} = \Gamma_i \cap \partial_{s_\lambda}(\Gamma_i). \quad (3.1)$$

The remainder of the set Γ_i we simply call the ‘interior’ and denote by Γ_i^{in} , where

$$\Gamma_i^{\text{in}} := \{x \in \Gamma_i : d_\infty(x, \partial\Gamma_i) > \alpha \lambda^{-1/d} \log \lambda\} = \Gamma_i \setminus \partial_{s_\lambda}(\Gamma_i).$$

As previously mentioned, we assume that ξ is translation invariant, and that $\mathcal{M} = \{1\}$.

Define

$$T_i^{\text{bd}} := \int_{\Gamma_i^{\text{bd}}} f_i d\mu_\lambda^\xi; \quad \text{and} \quad T_i^{\text{in}} := \int_{\Gamma_i^{\text{in}}} f_i d\mu_\lambda^\xi,$$

so that $T_i = T_i^{\text{in}} + T_i^{\text{bd}}$. To prepare for the proof of Theorem 2.1 we need some auxiliary lemmas. For the subsequent results, we will need the following covering of scaled-up Borel regions $\lambda^{1/d}B \subset \mathbb{R}^d$ by cubes of side 1.

First we need some more notation. Let $\text{card}(\mathcal{X})$ denote the cardinality of set \mathcal{X} . For $x \in \mathbb{R}^d$, let Q_x denote the unit-volume ℓ_∞ ball in \mathbb{R}^d with centre x (i.e., the unit d -cube at x). For a Borel set $B \subseteq A \subset \mathbb{R}^d$, let

$$\mathcal{Z}_\lambda(B) := \{x \in \mathbb{Z}^d : Q_x \cap \lambda^{1/d}B \neq \emptyset\}, \quad (3.2)$$

and set $n_\lambda(B) := \text{card}(\mathcal{Z}_\lambda(B))$. Then the covering of $\lambda^{1/d}B$ is

$$\mathcal{Q}_\lambda(B) := \{Q_z : z \in \mathcal{Z}_\lambda(B)\}. \quad (3.3)$$

The next result gives error bounds for approximating the volume of $\lambda^{1/d}\Gamma_i$ or of $\lambda^{1/d}\Gamma_i^{\text{bd}}$ (as defined at (3.1)) by the number of unit cubes in \mathbb{Z}^d in its covering (as defined at (3.2) and (3.3)).

Lemma 3.1 *Let Γ_i be a non-null Borel subset of $A \subset \mathbb{R}^d$ such that $|\partial_r(\Gamma_i)| = O(r)$ as $r \downarrow 0$. Then, as $\lambda \rightarrow \infty$,*

$$n_\lambda(\Gamma_i) - |\lambda^{1/d}\Gamma_i| = O(\lambda^{(d-1)/d}). \quad (3.4)$$

Define Γ_i^{bd} as at (3.1). Then, as $\lambda \rightarrow \infty$,

$$n_\lambda(\Gamma_i^{\text{bd}}) - |\lambda^{1/d}\Gamma_i^{\text{bd}}| = O(\lambda^{(d-1)/d} \log \lambda). \quad (3.5)$$

Proof. There exists a constant $c \in (0, \infty)$ (depending only on d) such that, for any $\lambda > 0$, and any non-null Borel subset B of A ,

$$\lambda^{1/d}B \subseteq \bigcup_{z \in \mathcal{Z}_\lambda(B)} Q_z \subseteq \lambda^{1/d}B \cup \partial_c(\lambda^{1/d}B),$$

and hence

$$|\lambda^{1/d}B| \leq n_\lambda(B) \leq |\lambda^{1/d}B| + |\partial_c(\lambda^{1/d}B)| = |\lambda^{1/d}B| + \lambda |\partial_{c\lambda^{-1/d}}(B)|. \quad (3.6)$$

In the case $B = \Gamma_i$, the regularity assumption that $|\partial_r(\Gamma_i)| = O(r)$ as $r \downarrow 0$ implies that $|\partial_{c\lambda^{-1/d}}(\Gamma_i)| = O(\lambda^{-1/d})$. Thus (3.4) follows from (3.6).

In the case $B = \Gamma_i^{\text{bd}}$, we have that

$$|\partial_{c\lambda^{-1/d}}(\Gamma_i^{\text{bd}})| \leq |\partial_{c\lambda^{-1/d}+s_\lambda}(\Gamma_i)| = O(s_\lambda),$$

as $\lambda \rightarrow \infty$, again by the regularity assumption on Γ_i . Thus (3.6) yields (3.5) in this case. \square

Once more consider a Borel subset B of $A \subset \mathbb{R}^d$ and the covering $\mathcal{Q}_\lambda(B)$ of $\lambda^{1/d}B$. For all $z \in \mathcal{Z}_\lambda(B)$, the number of points of $\mathcal{P}_\lambda \cap \lambda^{-1/d}Q_z$ is a Poisson random variable N_z with parameter $\nu_z := \lambda \int_{\lambda^{-1/d}Q_z} \kappa(x)dx$. Assuming $\nu_z > 0$, choose an ordering on the points of $\mathcal{P}_\lambda \cap \lambda^{-1/d}Q_z$ uniformly at random from all $N_z!$ possibilities. List the points as $X_{z,1}, \dots, X_{z,N_z}$, where conditional on the value of N_z , the random variables $X_{z,k}$, $k = 1, 2, \dots, N_z$ are i.i.d. on $\lambda^{-1/d}Q_z$ with a density $\kappa_z(\cdot) := \kappa(\cdot) / \int_{\lambda^{-1/d}Q_z} \kappa(x)dx$. Thus we have the representation

$$\mathcal{P}_\lambda \cap B = \bigcup_{z \in \mathcal{Z}_\lambda(B)} \bigcup_{k=1}^{N_z} (\{X_{z,k}\} \cap B).$$

Then for f in $\mathcal{B}(B)$, we can express $\langle f, \mu_\lambda^\xi \rangle$ as follows:

$$\langle f, \mu_\lambda^\xi \rangle = \sum_{z \in \mathcal{Z}_\lambda(B)} \sum_{k=1}^{N_z} \xi_\lambda(X_{z,k}; \mathcal{P}_\lambda) \cdot f(X_{z,k}) \cdot \mathbf{1}_B(X_{z,k}). \quad (3.7)$$

For all $z \in \mathcal{Z}_\lambda(B)$ and for all $k \in \mathbb{N}$, let $R_{z,k}$ denote the radius of stabilization of ξ at $X_{z,k}$ if $1 \leq k \leq N_z$ and let $R_{z,k} = 0$ otherwise. Define the event $E_{z,k} := \{R_{z,k} \leq \alpha \log \lambda\}$. We define here the function $\tilde{T}(B; f)$ as follows, the idea being that $\tilde{T}(B; f)$ is, with high probability, the same as $\langle f, \mu_\lambda^\xi \rangle$, but exhibits a much more localized dependency structure. Set

$$\tilde{T}(B; f) := \sum_{z \in \mathcal{Z}_\lambda(B)} \sum_{k=1}^{N_z} \xi_\lambda(X_{z,k}; \mathcal{P}_\lambda) \cdot \mathbf{1}_{E_{z,k}} \cdot f(X_{z,k}) \cdot \mathbf{1}_B(X_{z,k}), \quad (3.8)$$

where we use $\mathbf{1}_E$ to denote the indicator random variable of the event E .

Recall that Γ_i , $i = 1, 2, \dots, m$ are disjoint non-null Borel regions in $A \subset \mathbb{R}^d$ and $f_i \in \mathcal{B}(\Gamma_i)$ for $i = 1, 2, \dots, m$. Then for each i , $\tilde{T}(\Gamma_i; f_i)$ is defined by (3.8). In the same way as we use the abbreviations T_i , T_i^{bd} and T_i^{in} , we let $\tilde{T}_i := \tilde{T}(\Gamma_i; f_i)$, $\tilde{T}_i^{\text{bd}} := \tilde{T}(\Gamma_i^{\text{bd}}; f_i)$, and $\tilde{T}_i^{\text{in}} := \tilde{T}(\Gamma_i^{\text{in}}; f_i)$. Thus $\tilde{T}_i = \tilde{T}_i^{\text{bd}} + \tilde{T}_i^{\text{in}}$.

For $z \in \mathcal{Z}_\lambda(B)$ let $Y_z(B; f)$ be the contribution to $\tilde{T}(B; f)$ from the points in $\lambda^{-1/d}Q_z$, i.e.

$$Y_z(B; f) := \sum_{k=1}^{N_z} \xi_\lambda(X_{z,k}; \mathcal{P}_\lambda) \cdot \mathbf{1}_{E_{z,k}} \cdot f(X_{z,k}) \cdot \mathbf{1}_B(X_{z,k}), \quad (3.9)$$

so that $\tilde{T}(B; f) = \sum_{z \in \mathcal{Z}_\lambda(B)} Y_z(B; f)$.

Let A_λ , $\lambda \geq 1$ be a family of Borel subsets of $A \subset \mathbb{R}^d$. The next two results show that the moments condition (2.1) implies bounds on the moments of $Y_z(A_\lambda; f)$ for $f \in \mathcal{B}(A)$. When we come to apply the two lemmas below, we will be taking $A_\lambda = \Gamma_i$ or $A_\lambda = \Gamma_i^{\text{bd}}$.

Lemma 3.2 *Let A_λ , $\lambda \geq 1$ be a family of Borel subsets of $A \subset \mathbb{R}^d$. If (2.1) holds for some $p > 0$, then there is a constant $C \in (0, \infty)$ such that for all $\lambda \geq 1$, all $k \geq 1$ and $z \in \mathcal{Z}_\lambda(A_\lambda)$*

$$\mathbb{E}[|\xi_\lambda(X_{z,k}; \mathcal{P}_\lambda) \cdot \mathbf{1}_{A_\lambda}(X_{z,k}) \cdot \mathbf{1}_{\{k \leq N_z\}}|^p] \leq C. \quad (3.10)$$

Proof. It suffices to consider the case with $A_\lambda = A$ for all λ . The proof of the lemma closely follows that of Lemma 4.2 in [16], although our covering is somewhat different. In the notation of the proof of Lemma 4.2 of [16], we have $\rho_\lambda = 1$ and $\nu_i = \nu_{z_i} \equiv \lambda \int_{\lambda^{-1/d}Q_{z_i}} \kappa(x) dx \leq \|\kappa\|_\infty$, where we have written $\mathcal{Z}_\lambda(B) = \{z_1, \dots, z_{n_\lambda(B)}\}$. Then, following the argument in [16], we obtain (3.10). \square

Lemma 3.3 *Let A_λ , $\lambda \geq 1$, be a sequence of Borel subsets of $A \subset \mathbb{R}^d$, and suppose $f \in \mathcal{B}(A)$. If (2.1) holds for some $p > 1$, then for any $q \in (1, p)$ there is a constant $C \in (0, \infty)$ such that for all $\lambda \geq 1$ and all $z \in \mathcal{Z}_\lambda(A_\lambda)$*

$$\|Y_z(A_\lambda; f)\|_q^q \leq C. \quad (3.11)$$

Proof. The proof closely follows that of Lemma 4.3 in [16], again with ρ_λ there equal to 1 (and $\nu_i \leq \|\kappa\|_\infty$). Thus, with the use of Lemma 3.2 (and the boundedness of f), we obtain (3.11). \square

Lemma 3.4 Suppose that ξ is exponentially stabilizing and satisfies the moments condition (2.1) for some $p > 3$. Then there exists a constant $C \in (0, \infty)$ such that for all $\lambda \geq 2$

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P} \left[\frac{\tilde{T}_i - \mathbb{E}[\tilde{T}_i]}{(\text{Var}[\tilde{T}_i])^{1/2}} \leq t \right] - \Phi(t) \right| \leq C \lambda (\text{Var}[\tilde{T}_i])^{-3/2} (\log \lambda)^{3d}. \quad (3.12)$$

Moreover, (3.12) holds with \tilde{T}_i replaced by \tilde{T}_i^{in} everywhere.

Proof. The statement for \tilde{T}_i follows from equation (4.18) in [16] with $\rho_\lambda = O(\log \lambda)$, $q = 3$, and taking the A_λ of [16] to be Γ_i . In equation (4.18) of [16], T'_λ is the equivalent of our \tilde{T}_i , T_λ is our T_i , and S is our $(\tilde{T}_i - \mathbb{E}[\tilde{T}_i])(\text{Var}[\tilde{T}_i])^{-1/2}$. The statement for \tilde{T}_i^{in} follows in the same way, this time taking the A_λ of [16] to be Γ_i^{in} . \square

Lemma 3.5 Suppose that (2.1) holds for some $p > 2$. Then there exist constants $C_1, C_2, C_3 \in (0, \infty)$ such that, for all $\lambda \geq 2$,

$$\text{Var}[\tilde{T}_i^{\text{bd}}] \leq C_1 \lambda^{(d-1)/d} (\log \lambda)^{d+1}, \quad (3.13)$$

$$\text{Var}[\tilde{T}_i] \leq C_2 \lambda (\log \lambda)^d, \quad \text{and} \quad (3.14)$$

$$\text{Var}[\tilde{T}_i^{\text{in}}] \leq C_3 \lambda (\log \lambda)^d. \quad (3.15)$$

Proof. First we prove (3.13). Consider the covering $\mathcal{Q}_\lambda(\Gamma_i^{\text{bd}})$ of $\lambda^{1/d} \Gamma_i^{\text{bd}}$ by unit d -cubes, as defined at (3.3). For $z \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})$ let $Y_z(\Gamma_i^{\text{bd}}; f_i)$ be the contribution to \tilde{T}_i^{bd} from the points in $\lambda^{-1/d} Q_z$, as defined at (3.9), that is

$$Y_z(\Gamma_i^{\text{bd}}; f_i) := \sum_{k=1}^{N_z} \xi_\lambda(X_{z,k}; \mathcal{P}_\lambda) \cdot \mathbf{1}_{E_{z,k}} \cdot f_i(X_{z,k}) \cdot \mathbf{1}_{\Gamma_i^{\text{bd}}}(X_{z,k}). \quad (3.16)$$

Now, using the representation $\tilde{T}_i^{\text{bd}} = \sum_{z \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})} Y_z(\Gamma_i^{\text{bd}}; f_i)$, we have

$$\text{Var}[\tilde{T}_i^{\text{bd}}] = \sum_z \text{Var}[Y_z(\Gamma_i^{\text{bd}}; f_i)] + \sum_{z \neq w} \text{Cov}[Y_z(\Gamma_i^{\text{bd}}; f_i), Y_w(\Gamma_i^{\text{bd}}; f_i)]. \quad (3.17)$$

By the assumption that (2.1) holds for some $p > 2$, by taking $q = 2$ and $A_\lambda = \Gamma_i^{\text{bd}}$ in Lemma 3.3 we have that $\text{Var}[Y_z(\Gamma_i^{\text{bd}}; f_i)] \leq V$, for some constant $V < \infty$, for all $z \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})$. So by the Cauchy-Schwarz inequality we have $\text{Cov}[Y_z(\Gamma_i^{\text{bd}}; f_i), Y_w(\Gamma_i^{\text{bd}}; f_i)] \leq V$. Also, $Y_z(\Gamma_i^{\text{bd}}; f_i)$ and $Y_w(\Gamma_i^{\text{bd}}; f_i)$ are independent if $d_2(Q_z, Q_w) > 2\alpha \log \lambda$ (by the definition of $E_{z,k}$). Further, given z , the number of w for which $d_2(Q_z, Q_w) \leq 2\alpha \log \lambda$ is $O((\log \lambda)^d)$. Hence (3.17) implies that

$$\text{Var}[\tilde{T}_i^{\text{bd}}] \leq n_\lambda(\Gamma_i^{\text{bd}})(V + O((\log \lambda)^d)). \quad (3.18)$$

Then by (3.5) we have that

$$n_\lambda(\Gamma_i^{\text{bd}}) = \lambda |\Gamma_i^{\text{bd}}| + O(\lambda^{(d-1)/d} \log \lambda) = O(\lambda^{(d-1)/d} \log \lambda), \quad (3.19)$$

using (3.1) and (A2). So from (3.18) and (3.19) we obtain (3.13).

The proof of (3.14) follows similarly, using $A_\lambda = \Gamma_i$ for all λ in Lemma 3.3 and (3.4) in place of (3.5). Finally, (3.15) follows from (3.14), (3.13) and the Cauchy-Schwarz inequality, since $\tilde{T}_i^{\text{in}} = \tilde{T}_i - \tilde{T}_i^{\text{bd}}$. \square

Lemma 3.6 *Suppose that ξ is exponentially stabilizing and satisfies the moments condition (2.1) for some $p > 3$. Then there exists a constant $C \in (0, \infty)$ such that for any $\delta > 0$, all $\lambda \geq 2$, and any $t \in \mathbb{R}$*

$$\mathbb{P} \left[\left| \frac{\tilde{T}_i - \mathbb{E}[\tilde{T}_i]}{(\text{Var}[\tilde{T}_i])^{1/2}} - t \right| \leq \delta \right] \leq \sqrt{\frac{2}{\pi}} \delta + C(\log \lambda)^{3d} \lambda (\text{Var}[\tilde{T}_i])^{-3/2}, \quad (3.20)$$

and also

$$\begin{aligned} \mathbb{P} \left[\left| \frac{\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]}{(\text{Var}[\tilde{T}_i])^{1/2}} - t \right| \leq \delta \right] &\leq 2\sqrt{\frac{2}{\pi}} \delta + C(\log \lambda)^{3d} \lambda (\text{Var}[\tilde{T}_i])^{-3/2} \\ &\quad + \mathbb{P} \left[\left| \frac{\tilde{T}_i^{\text{bd}} - \mathbb{E}[\tilde{T}_i^{\text{bd}}]}{(\text{Var}[\tilde{T}_i])^{1/2}} \right| > \delta \right]. \end{aligned} \quad (3.21)$$

Proof. First we prove (3.20). For the duration of this proof, write

$$F(t) = \mathbb{P} \left[\frac{\tilde{T}_i - \mathbb{E}[\tilde{T}_i]}{(\text{Var}[\tilde{T}_i])^{1/2}} \leq t \right].$$

Then we have that for $t \in \mathbb{R}$ and $\delta > 0$

$$\begin{aligned} \mathbb{P} \left[\left| \frac{\tilde{T}_i - \mathbb{E}[\tilde{T}_i]}{(\text{Var}[\tilde{T}_i])^{1/2}} - t \right| \leq \delta \right] &= F(t + \delta) - F(t - \delta) \\ &= \Phi(t + \delta) - \Phi(t - \delta) + [F(t + \delta) - \Phi(t + \delta)] - [F(t - \delta) - \Phi(t - \delta)] \\ &\leq |\Phi(t + \delta) - \Phi(t - \delta)| + |F(t + \delta) - \Phi(t + \delta)| + |F(t - \delta) - \Phi(t - \delta)|. \end{aligned}$$

Then (3.20) follows from the Mean Value Theorem (applied to the first term on the right of the above inequality) and Lemma 3.4 (applied to the other two terms). Finally, we have that for $\delta > 0$

$$\mathbb{P} \left[\left| \frac{\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]}{(\text{Var}[\tilde{T}_i])^{1/2}} - t \right| \leq \delta \right] \leq \mathbb{P} \left[\left| \frac{\tilde{T}_i - \mathbb{E}[\tilde{T}_i]}{(\text{Var}[\tilde{T}_i])^{1/2}} - t \right| \leq 2\delta \right] + \mathbb{P} \left[\left| \frac{\tilde{T}_i^{\text{bd}} - \mathbb{E}[\tilde{T}_i^{\text{bd}}]}{(\text{Var}[\tilde{T}_i])^{1/2}} \right| > \delta \right].$$

Then using (3.20) yields (3.21). \square

Lemma 3.7 *Suppose that the moments condition (2.1) holds for all $p \geq 1$, and condition (A2) holds. Let k be an even positive integer. Then there exists a constant $C \in (0, \infty)$ (depending on k) such that for all $\lambda \geq 2$,*

$$\mathbb{E} \left[\left| \tilde{T}_i^{\text{bd}} - \mathbb{E}[\tilde{T}_i^{\text{bd}}] \right|^k \right] \leq C \lambda^{k(d-1)/(2d)} (\log \lambda)^{k(1+d)/2}. \quad (3.22)$$

Proof. Again consider the covering $\mathcal{Q}_\lambda(\Gamma_i^{\text{bd}})$ of $\lambda^{1/d} \Gamma_i^{\text{bd}}$ as defined at (3.3). For $z \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})$, let \bar{Y}_z be the contribution to $\tilde{T}_i^{\text{bd}} - \mathbb{E}[\tilde{T}_i^{\text{bd}}]$ from cube Q_z , that is $\bar{Y}_z := Y_z(\Gamma_i^{\text{bd}}; f_i) - \mathbb{E}[Y_z(\Gamma_i^{\text{bd}}; f_i)]$ where $Y_z(\Gamma_i^{\text{bd}}; f_i)$ is given by (3.16). Thus, for all $z \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})$, $\mathbb{E}[\bar{Y}_z] = 0$ and $\text{Var}[\bar{Y}_z] \leq V$ for constant V , by Lemma 3.3.

Let k be an even positive integer. Then

$$\mathbb{E} \left[\left| \tilde{T}_i^{\text{bd}} - \mathbb{E} \tilde{T}_i^{\text{bd}} \right|^k \right] = \sum_{z_1 \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})} \sum_{z_2 \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})} \cdots \sum_{z_k \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})} \mathbb{E} [\bar{Y}_{z_1} \bar{Y}_{z_2} \cdots \bar{Y}_{z_k}].$$

The term $\mathbb{E} [\bar{Y}_{z_1} \bar{Y}_{z_2} \cdots \bar{Y}_{z_k}]$ will vanish if any of the cubes corresponding to the \bar{Y}_{z_j} is farther than $2\alpha \log \lambda$ from all the other cubes (since then it will be independent of the other \bar{Y}_{z_j} and has expectation zero). In other words, the term vanishes if the appropriate geometric graph (in the sense of [9]) on z_1, z_2, \dots, z_k has any isolated vertices. For a non-zero contribution to the sum, we require the graph to have no isolated vertices — so it must have no more than $k/2$ components. So in effect, there are at most $k/2$ ‘free’ indices of (z_1, \dots, z_k) . Values that are not ‘free’ have $O((\log \lambda)^d)$ possible values.

Further, $\mathbb{E} [\bar{Y}_{z_1} \bar{Y}_{z_2} \cdots \bar{Y}_{z_k}] \leq C$ for some constant C , by Lemma 3.3 (given the moments condition (2.1) for all $p \geq 1$) and Hölder’s inequality. Thus for some other constant also denoted C ,

$$\begin{aligned} \sum_{z_1 \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})} \sum_{z_2 \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})} \cdots \sum_{z_k \in \mathcal{Z}_\lambda(\Gamma_i^{\text{bd}})} \mathbb{E} [\bar{Y}_{z_1} \bar{Y}_{z_2} \cdots \bar{Y}_{z_k}] &\leq C(n_\lambda(\Gamma_i^{\text{bd}}))^{k/2} (\log \lambda)^{kd/2} \\ &\leq C\lambda^{k(d-1)/(2d)} (\log \lambda)^{k/2} (\log \lambda)^{kd/2}, \end{aligned}$$

the final inequality by (3.5), (3.1) and (A2). Hence we have (3.22). \square

The next lemma says that given condition (A1), we can obtain lower bounds on the variances of \tilde{T}_i^{in} and \tilde{T}_i . We will need the following result from [16] (see (4.17) therein), which says that if ξ is exponentially stabilizing and satisfies the moments condition (2.1) for some $p > 2$, then

$$\left| \text{Var}[\tilde{T}_i] - \text{Var}[T_i] \right| \leq C\lambda^{-2}. \quad (3.23)$$

Lemma 3.8 *Suppose that (A1) and (A2) are satisfied, and that the moments condition (2.1) holds for all $p \geq 1$. Then there exist constants $C \in (0, \infty)$ and $\lambda_0 \in [1, \infty)$ such that for all $\lambda \geq \lambda_0$*

$$\text{Var}[\tilde{T}_i] \geq C\lambda, \quad \text{and} \quad (3.24)$$

$$\text{Var}[\tilde{T}_i^{\text{in}}] \geq C\lambda. \quad (3.25)$$

Proof. These follow in a straightforward manner from (3.23), (A1), (3.13), (3.14) and the Cauchy-Schwarz inequality. \square

Lemma 3.9 *Suppose that ξ is exponentially stabilizing and satisfies the moments condition (2.1) for all $p \geq 1$. Suppose conditions (A1) and (A2) hold. Then for any $\varepsilon > 0$, there exists $C \in (0, \infty)$ such that for all $\lambda \geq 1$*

$$\begin{aligned} &\left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ \frac{\tilde{T}_i - \mathbb{E}[\tilde{T}_i]}{(\text{Var}[\tilde{T}_i])^{1/2}} \leq t_i \right\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \\ &\leq C\lambda^{-1/(2d+\varepsilon)} + \left| \prod_{i=1}^m \mathbb{P} \left[\frac{\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]}{(\text{Var}[\tilde{T}_i])^{1/2}} \leq t_i \right] - \prod_{i=1}^m \Phi(t_i) \right|. \end{aligned} \quad (3.26)$$

Proof. We abbreviate our notation for the duration of the current proof by setting $\sigma_i := (\text{Var}[\tilde{T}_i])^{1/2}$. Then we have

$$\begin{aligned}
& \left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ (\tilde{T}_i - \mathbb{E}[\tilde{T}_i]) \sigma_i^{-1} \leq t_i \right\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \\
& \leq \left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ (\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} \leq t_i \right\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \\
& \quad + \sum_{i=1}^m \mathbb{P} \left[(\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} \leq t_i, (\tilde{T}_i - \mathbb{E}[\tilde{T}_i]) \sigma_i^{-1} > t_i \right] \\
& \quad + \sum_{i=1}^m \mathbb{P} \left[(\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} > t_i, (\tilde{T}_i - \mathbb{E}[\tilde{T}_i]) \sigma_i^{-1} \leq t_i \right]. \tag{3.27}
\end{aligned}$$

For any $\beta > 0$, we have

$$\begin{aligned}
& \max \left(\mathbb{P} \left[(\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} \leq t, (\tilde{T}_i - \mathbb{E}[\tilde{T}_i]) \sigma_i^{-1} > t \right], \right. \\
& \quad \left. \mathbb{P} \left[(\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} > t, (\tilde{T}_i - \mathbb{E}[\tilde{T}_i]) \sigma_i^{-1} \leq t \right] \right) \\
& \leq \mathbb{P} \left[\left| (\tilde{T}_i^{\text{bd}} - \mathbb{E}[\tilde{T}_i^{\text{bd}}]) \sigma_i^{-1} \right| > \lambda^{-\beta} \right] + \mathbb{P} \left[\left| (\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} - t \right| \leq \lambda^{-\beta} \right]. \tag{3.28}
\end{aligned}$$

Then, from (3.27) and (3.28)

$$\begin{aligned}
& \left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ (\tilde{T}_i - \mathbb{E}[\tilde{T}_i]) \sigma_i^{-1} \leq t_i \right\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \\
& \leq \left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ (\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} \leq t_i \right\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \\
& + 2 \sum_{i=1}^m \mathbb{P} \left[\left| (\tilde{T}_i^{\text{bd}} - \mathbb{E}[\tilde{T}_i^{\text{bd}}]) \sigma_i^{-1} \right| > \lambda^{-\beta} \right] + 2 \sum_{i=1}^m \mathbb{P} \left[\left| (\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} - t_i \right| \leq \lambda^{-\beta} \right]. \tag{3.29}
\end{aligned}$$

Since $d_2(\lambda^{1/d} \Gamma_i^{\text{in}}, \lambda^{1/d} \Gamma_j^{\text{in}})$ is at least $2\alpha \log \lambda$ for $i \neq j$, $\tilde{T}_i^{\text{in}}, 1 \leq i \leq m$ is a sequence of mutually independent random variables, so that

$$\mathbb{P} \left[\bigcap_{i=1}^m \left\{ (\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} \leq t_i \right\} \right] = \prod_{i=1}^m \mathbb{P} \left[(\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]) \sigma_i^{-1} \leq t_i \right]. \tag{3.30}$$

Also, from Markov's inequality, we have that, for $k \in 2\mathbb{N}$,

$$\mathbb{P} \left[\left| (\tilde{T}_i^{\text{bd}} - \mathbb{E}[\tilde{T}_i^{\text{bd}}]) \sigma_i^{-1} \right| > \lambda^{-\beta} \right] \leq \mathbb{E} \left[\left| \tilde{T}_i^{\text{bd}} - \mathbb{E}[\tilde{T}_i^{\text{bd}}] \right|^k \right] \left(\text{Var}[\tilde{T}_i] \right)^{-k/2} \lambda^{k\beta}. \tag{3.31}$$

Then we obtain, from (3.31), with (3.22) and (3.24),

$$\mathbb{P} \left[\left| (\tilde{T}_i^{\text{bd}} - \mathbb{E}[\tilde{T}_i^{\text{bd}}]) \sigma_i^{-1} \right| > \lambda^{-\beta} \right] \leq C \lambda^{k(\beta-1/(2d))} (\log \lambda)^{k(1+d)/2}; \tag{3.32}$$

this then gives a bound for the penultimate sum in (3.29). To bound the final sum in (3.29), taking $\delta = \lambda^{-\beta}$ we have from (3.21), (3.32) and (3.24) that

$$\begin{aligned} & \mathbb{P} \left[\left| (\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}])\sigma_i^{-1} - t_i \right| \leq \lambda^{-\beta} \right] \\ & \leq 2\sqrt{\frac{2}{\pi}}\lambda^{-\beta} + C(\log \lambda)^{3d}\lambda^{-1/2} + C\lambda^{k(\beta-1/(2d))}(\log \lambda)^{k(1+d)/2}. \end{aligned} \quad (3.33)$$

To obtain the best rates of convergence via this method, we want to maximize the lowest power of λ^{-1} on the right-hand sides of (3.32) and (3.33). So we choose β such that $-\beta = k(\beta - 1/(2d))$, that is, take

$$\beta = \frac{k}{2d(k+1)}. \quad (3.34)$$

For any $\varepsilon > 0$ we can choose k large enough in (3.34) to give $1/(2d) > \beta \geq 1/(2d + \varepsilon/2)$. Then, for λ sufficiently large, $\lambda^{-1/(2d+\varepsilon)} \geq \lambda^{-1/(2d+\varepsilon/2)}(\log \lambda)^{k(1+d)/2}$. Now from (3.29) and (3.30), with the bounds (3.32) and (3.33) we obtain (3.26). This completes the proof of the lemma. \square

Proof of Theorem 2.1. To complete the proof we proceed in a similar manner to [16]. Let

$$E_\lambda := \bigcap_{i=1}^m \bigcap_{z \in Z_\lambda(\Gamma_i)} \bigcap_{k=1}^{N_z} E_{z,k},$$

recalling the definition of the event $E_{z,k}$ just below (3.7). By standard Palm theory (e.g. Theorem 1.6 in [9]) and exponential stabilization (see (4.11) in [16]), we have that $\mathbb{P}[E_\lambda^c] \leq C\lambda^{-3}$ for λ sufficiently large and some $C \in (0, \infty)$. Then $|\tilde{T}_i - T_i| = 0$ except possibly on the set E_λ^c , which has probability less than $C\lambda^{-3}$.

For $i = 1, \dots, m$, let $K_i := (\text{Var}[\tilde{T}_i])^{-1/2}(\tilde{T}_i - \mathbb{E}[\tilde{T}_i])$ and $Z_i := (\text{Var}[\tilde{T}_i])^{-1/2}(T_i - \mathbb{E}[T_i])$. Then for $\delta > 0$ we have that for any $t_i \in \mathbb{R}$

$$\{(Z_i \leq t_i)\Delta(K_i \leq t_i)\} \subseteq \{|K_i - t_i| \leq \delta\} \cup \{|Z_i - K_i| \geq \delta\},$$

so that

$$\begin{aligned} & \left| \mathbb{P} \left[\bigcap_{i=1}^m \{Z_i \leq t_i\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \leq \left| \mathbb{P} \left[\bigcap_{i=1}^m \{K_i \leq t_i\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \\ & + \sum_{i=1}^m \mathbb{P}[|K_i - t_i| \leq \delta] + \sum_{i=1}^m \mathbb{P}[|Z_i - K_i| \geq \delta]. \end{aligned} \quad (3.35)$$

Then, using (3.26) for the first term on the right-hand side of the inequality in (3.35), and (3.20) with (3.24) for the second, we obtain

$$\begin{aligned} & \left| \mathbb{P} \left[\bigcap_{i=1}^m \{Z_i \leq t_i\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \leq C\lambda^{-1/(2d+\varepsilon)} + \left| \prod_{i=1}^m \mathbb{P} \left[\frac{\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]}{(\text{Var}[\tilde{T}_i])^{1/2}} \leq t_i \right] - \prod_{i=1}^m \Phi(t_i) \right| \\ & + C\delta + C(\log \lambda)^{3d}\lambda^{-1/2} + \sum_{i=1}^m \mathbb{P}[|Z_i - K_i| \geq \delta] \end{aligned} \quad (3.36)$$

We now consider the second term on the right-hand side of (3.36). For ease of notation, write

$$G_i(t) := \mathbb{P} \left[\frac{\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]}{(\text{Var}[\tilde{T}_i^{\text{in}}])^{1/2}} \leq t \right],$$

for $i = 1, \dots, m$. Then

$$\begin{aligned} \left| \prod_{i=1}^m G_i(t_i) - \prod_{i=1}^m \Phi(t_i) \right| &= \left| \sum_{i=1}^m [G_i(t_i) - \Phi(t_i)] \prod_{j=i+1}^m \Phi(t_j) \prod_{k=1}^{i-1} G_k(t_k) \right| \\ &\leq \sum_{i=1}^m |G_i(t_i) - \Phi(t_i)|. \end{aligned} \quad (3.37)$$

Writing

$$H_i(t) := \mathbb{P} \left[\frac{\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]}{(\text{Var}[\tilde{T}_i^{\text{in}}])^{1/2}} \leq t \right],$$

we have that, for $i = 1, \dots, m$

$$|G_i(t_i) - \Phi(t_i)| \leq |H_i(t_i(1 + \gamma_i)) - \Phi(t_i(1 + \gamma_i))| + |\Phi(t_i(1 + \gamma_i)) - \Phi(t_i)|, \quad (3.38)$$

where $1 + \gamma_i := \left(\frac{\text{Var}[\tilde{T}_i]}{\text{Var}[\tilde{T}_i^{\text{in}}]} \right)^{1/2}$. Then, using Lemma 3.4 we have that the first term on the right-hand side of (3.38) satisfies

$$|H_i(t_i(1 + \gamma_i)) - \Phi(t_i(1 + \gamma_i))| \leq C(\log \lambda)^{3d} \lambda (\text{Var}[\tilde{T}_i^{\text{in}}])^{-3/2} \leq C\lambda^{-1/2} (\log \lambda)^{3d}, \quad (3.39)$$

by (3.25). In order to deal with the second term on the right-hand side of (3.38), we need to estimate γ_i . We note that

$$\frac{\text{Var}[\tilde{T}_i]}{\text{Var}[\tilde{T}_i^{\text{in}}]} = 1 + \frac{\text{Var}[\tilde{T}_i^{\text{bd}}]}{\text{Var}[\tilde{T}_i^{\text{in}}]} + \frac{2\text{Cov}[\tilde{T}_i^{\text{in}}, \tilde{T}_i^{\text{bd}}]}{\text{Var}[\tilde{T}_i^{\text{in}}]}.$$

Then using the upper and lower variance bounds (3.13), (3.15), (3.25), and the Cauchy-Schwarz inequality, yields

$$\frac{\text{Var}[\tilde{T}_i]}{\text{Var}[\tilde{T}_i^{\text{in}}]} = 1 + O(\lambda^{-1/(2d)} (\log \lambda)^{(2d+1)/2}),$$

so that

$$\gamma_i = O(\lambda^{-1/(2d)} (\log \lambda)^{(2d+1)/2}). \quad (3.40)$$

Since for all $s \leq t$ we have $|\Phi(s) - \Phi(t)| \leq (t - s) \sup_{s \leq u \leq t} \varphi(u)$ (where φ is the standard normal density function), we have

$$\begin{aligned} &\sup_{t_i} |\Phi(t_i(1 + \gamma_i)) - \Phi(t_i)| \\ &\leq C \sup_{t_i} \left(|t_i| \lambda^{-1/(2d)} (\log \lambda)^{(2d+1)/2} \sup_{|u - t_i| \leq t_i C \lambda^{-1/(2d)} (\log \lambda)^{(2d+1)/2}} \varphi(u) \right) \\ &\leq C \lambda^{-1/(2d)} (\log \lambda)^{(2d+1)/2}. \end{aligned} \quad (3.41)$$

So, for the second term on the right-hand side in (3.36), we obtain from (3.37), (3.38), (3.39) and (3.41)

$$\begin{aligned} & \sup_{t_1, \dots, t_m} \left| \prod_{i=1}^m \mathbb{P} \left[\frac{\tilde{T}_i^{\text{in}} - \mathbb{E}[\tilde{T}_i^{\text{in}}]}{(\text{Var}[\tilde{T}_i])^{1/2}} \leq t_i \right] - \prod_{i=1}^m \Phi(t_i) \right| \\ & \leq C(\log \lambda)^{3d} \lambda^{-1/2} + C\lambda^{-1/(2d)} (\log \lambda)^{(2d+1)/2}. \end{aligned} \quad (3.42)$$

We now move on to the fifth term on the right-hand side of (3.36). We have

$$|Z_i - K_i| = (\text{Var}[\tilde{T}_i])^{-1/2} |(T_i - \mathbb{E}[T_i]) - (\tilde{T}_i - \mathbb{E}[\tilde{T}_i])| \leq (\text{Var}[\tilde{T}_i])^{-1/2} (|T_i - \tilde{T}_i| + \mathbb{E}[|T_i - \tilde{T}_i|]),$$

and from just below (4.19) in [16], we have that this is bounded by $C\lambda^{-3}$ except possibly on the set E_λ^c which has probability less than $C\lambda^{-3}$. Thus by (3.36) with $\delta = C\lambda^{-3}$, and using (3.42) for the second term on the right-hand side of (3.36), we obtain

$$\begin{aligned} & \sup_{t_1, \dots, t_m} \left| \mathbb{P} \left[\bigcap_{i=1}^m \{Z_i \leq t_i\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \leq C\lambda^{-1/(2d+\varepsilon)} + C(\log \lambda)^{3d} \lambda^{-1/2} \\ & \quad + C\lambda^{-1/(2d)} (\log \lambda)^{(2d+1)/2} + C\lambda^{-3} = O(\lambda^{-1/(2d+\varepsilon)}). \end{aligned} \quad (3.43)$$

By the triangle inequality we have

$$\begin{aligned} & \sup_{t_1, \dots, t_m} \left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ \frac{T_i - \mathbb{E}[T_i]}{(\text{Var}[T_i])^{1/2}} \leq t_i \right\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \\ & \leq \sup_{t_1, \dots, t_m} \left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ \frac{T_i - \mathbb{E}[T_i]}{(\text{Var}[\tilde{T}_i])^{1/2}} \leq t_i \cdot \left(\frac{\text{Var}[T_i]}{\text{Var}[\tilde{T}_i]} \right)^{1/2} \right\} \right] - \prod_{i=1}^m \Phi \left(t_i \cdot \left(\frac{\text{Var}[T_i]}{\text{Var}[\tilde{T}_i]} \right)^{1/2} \right) \right| \\ & \quad + \sup_{t_1, \dots, t_m} \left| \prod_{i=1}^m \Phi \left(t_i \cdot \left(\frac{\text{Var}[T_i]}{\text{Var}[\tilde{T}_i]} \right)^{1/2} \right) - \prod_{i=1}^m \Phi(t_i) \right|. \end{aligned} \quad (3.44)$$

Now from (3.23) and (3.24), there is a constant $C \in (0, \infty)$ such that for all $\lambda \geq 1$ and all $t_i \in \mathbb{R}$

$$\begin{aligned} & \left| t_i \cdot \left(\frac{\text{Var}[T_i]}{\text{Var}[\tilde{T}_i]} \right)^{1/2} - t_i \right| = |t_i| \left| \left(1 + \frac{\text{Var}[T_i] - \text{Var}[\tilde{T}_i]}{\text{Var}[\tilde{T}_i]} \right)^{1/2} - 1 \right| \\ & = |t_i| \left| (1 + O(\lambda^{-3}))^{1/2} - 1 \right| \leq C|t_i|\lambda^{-3}; \end{aligned}$$

then since for all $s \leq t$ we have $|\Phi(s) - \Phi(t)| \leq (t - s) \max_{s \leq u \leq t} \varphi(u)$, we get

$$\sup_{t_i} \left| \Phi \left(t_i \cdot \left(\frac{\text{Var}[T_i]}{\text{Var}[\tilde{T}_i]} \right)^{1/2} \right) - \Phi(t_i) \right| \leq C \sup_{t_i} \left(|t_i|\lambda^{-3} \sup_{u: |u-t_i| \leq t_i C\lambda^{-3}} \varphi(u) \right) \leq C\lambda^{-3} \quad (3.45)$$

Then, considering the second term on the right-hand side of (3.44), we have

$$\sup_{t_1, \dots, t_m} \left| \prod_{i=1}^m \Phi \left(t_i \cdot \left(\frac{\text{Var}[T_i]}{\text{Var}[\tilde{T}_i]} \right)^{1/2} \right) - \prod_{i=1}^m \Phi(t_i) \right|$$

$$\begin{aligned}
&= \sup_{t_1, \dots, t_m} \left| \sum_{i=1}^m \left[\Phi \left(t_i \cdot \left(\frac{\text{Var}[T_i]}{\text{Var}[\tilde{T}_i]} \right)^{1/2} \right) - \Phi(t_i) \right] \prod_{j=i+1}^m \Phi(t_j) \prod_{k=1}^{i-1} \Phi \left(t_k \cdot \left(\frac{\text{Var}[T_k]}{\text{Var}[\tilde{T}_k]} \right)^{1/2} \right) \right| \\
&\leq \sup_{t_1, \dots, t_m} \sum_{i=1}^m \left| \Phi \left(t_i \cdot \left(\frac{\text{Var}[T_i]}{\text{Var}[\tilde{T}_i]} \right)^{1/2} \right) - \Phi(t_i) \right| \leq C\lambda^{-3},
\end{aligned} \tag{3.46}$$

by (3.45). Thus for any $\varepsilon > 0$, from (3.44) and (3.43) with (3.46),

$$\sup_{t_1, \dots, t_m} \left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ \frac{T_i - \mathbb{E}[T_i]}{(\text{Var}[T_i])^{1/2}} \leq t_i \right\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \leq C\lambda^{-1/(2d+\varepsilon)} + C\lambda^{-3} = O(\lambda^{-1/(2d+\varepsilon)}).$$

This completes the proof of Theorem 2.1. \square

4 Indication of applications

In applying Theorem 2.1, one needs to check that the stabilization and moments conditions given in Definitions 2.1 and 2.2 hold. These conditions, or related versions thereof, are known to hold for many problems of interest in geometric probability; see [4] and [16] for an indication of problems for which exponential stabilization and moment bounds are satisfied.

One also needs to verify the variance bound (A1): we discuss methods of doing this in Section 4.1 below. In many cases, (A1) (or related versions thereof) has been demonstrated, see for example [2, 4, 13].

In Section 4.2 we give an example of our result as applied to the k -nearest neighbour graph. In particular, we give a multivariate CLT with explicit variance scalings in the case of the nearest-neighbour (directed) graph on disjoint subsets of the real line (Theorem 4.1 below).

4.1 Control of variances

In this section we discuss conditions under which one can say something about the variances $\text{Var}[T_i]$. Recall that Theorem 2.1 is stated under assumption (A1). First we give a sufficient condition for (A1) to hold, similar in spirit to that used by Avram and Bertsimas [2]. Once again, for notational convenience we consider only the unmarked case with $\mathcal{M} = \{1\}$.

First we introduce some notation. Recall that Q_x denotes the unit d -cube centred at $x \in \mathbb{R}^d$. For a non-null Borel subset B of $A \subset \mathbb{R}^d$ and $\lambda > 0$ we define the following packing of $\lambda^{1/d}B$ by unit d -cubes. For $\lambda^{1/d}B \subset \mathbb{R}^d$ let

$$\mathcal{W}_\lambda(B) := \{w \in \mathbb{Z}^d : Q_w \subseteq \lambda^{1/d}B\}, \tag{4.1}$$

and set $m_\lambda(B) := \text{card}(\mathcal{W}_\lambda(B))$. Then we define the packing $\mathcal{K}_\lambda(B)$ of $\lambda^{1/d}B$ by

$$\mathcal{K}_\lambda(B) := \{Q_w : w \in \mathcal{W}_\lambda(B)\}. \tag{4.2}$$

Let $f \in \mathcal{B}(B)$. For $w \in \mathcal{W}_\lambda(B)$ set

$$F_w := F_\lambda(Q_w; B) := \sum_{x \in \mathcal{P}_\lambda \cap \lambda^{-1/d}Q_w} \xi_\lambda(x; \mathcal{P}_\lambda) \cdot f(x). \tag{4.3}$$

Let \mathcal{F}_λ denote the σ -field generated by the points of \mathcal{P}_λ .

Definition 4.1 Let $\{A_w : w \in \mathcal{W}_\lambda(B)\}$ be a set of $m_\lambda(B)$ events in \mathcal{F}_λ , associated with the $m_\lambda(B)$ cubes $Q_w(B)$, so that each A_w occurs with probability uniformly bounded away from zero. Let $J = \{w_1, \dots, w_M\}$ be the (random) set of indices $w \in \mathcal{W}_\lambda(B)$ such that A_w occurs. Let \mathcal{G}_λ denote the σ -field generated by the random set $J \subseteq \mathcal{W}_\lambda(B)$ and the values of $F_\lambda(Q_w; B)$ for $w \notin J$.

We say that μ_λ^ξ is nondegenerate on (B, f) if there exist events $\{A_w : w \in \mathcal{W}_\lambda(B)\}$ in \mathcal{F}_λ , with $\mathbb{P}(A_w) \geq \rho > 0$ for all w , such that:

- (i) given \mathcal{G}_λ , for all $w \in J$, $\text{Var}[F_w | \mathcal{G}_\lambda] > \eta > 0$ a.s.;
- (ii) given \mathcal{G}_λ , for all $w, v \in J$ with $w \neq v$, F_w and F_v are (conditionally) independent.

The idea of this condition is that the events A_w essentially ‘isolate’ cubes Q_w , while allowing strictly positive variability (of the integrated measure) within the cube, and a positive fraction of all the cubes Q_w will be so ‘isolated’.

This nondegeneracy condition can often be demonstrated. In many cases, event A_w will involve a configuration of many points in an ‘annulus’ just outside the cube Q_w , and an empty ‘moat’ inside the cube, that ensures sufficient independence; see [2] for such a construction (in a similar context) for the total length of the j -th nearest-neighbour, Voronoi, and Delaunay graphs.

We now show that given the nondegeneracy condition of Definition 4.1, we have lower bounds of order λ on the variances of T_i .

Lemma 4.1 Let Γ be a non-null Borel subset of $A \subset \mathbb{R}^d$ such that $|\partial_r(\Gamma)| = O(r)$ as $r \downarrow 0$. Then, as $\lambda \rightarrow \infty$,

$$m_\lambda(\Gamma) - |\lambda^{1/d}\Gamma| = O(\lambda^{(d-1)/d}). \quad (4.4)$$

Proof. This follows in a similar way to the proof of (3.4) given previously. \square

Lemma 4.2 Suppose that μ_λ^ξ is nondegenerate on (Γ_i, f_i) (see Definition 4.1). Then there exists a constant $C_i \in (0, \infty)$ such that for all λ sufficiently large

$$\text{Var}[T_i] \geq C_i \lambda.$$

Proof. From the definitions of the packing and covering defined by (4.1), (4.2) and (3.2), (3.3) respectively, and the equations (3.7) and (4.3), we have that for Borel $\Gamma \subseteq A \subset \mathbb{R}^d$ and $f \in \mathcal{B}(\Gamma)$

$$\langle f, \mu_\lambda^\xi \rangle = \sum_{w \in \mathcal{W}_\lambda(\Gamma)} F_\lambda(Q_w; \Gamma) + \Delta_\lambda(\Gamma), \quad (4.5)$$

where we have set

$$\Delta_\lambda(\Gamma) := \sum_{w \in \mathcal{Z}_\lambda(\Gamma) \setminus \mathcal{W}_\lambda(\Gamma)} \sum_{k=1}^{N_w} \xi_\lambda(X_{w,k}; \mathcal{P}_\lambda) \cdot f(X_{w,k}) \cdot \mathbf{1}_\Gamma(X_{w,k}).$$

That is, $\Delta_\lambda(\Gamma)$ gives the contributions to $\langle f, \mu_\lambda^\xi \rangle$ from cubes that are in the covering of $\lambda^{1/d}\Gamma$ but not the packing.

Consider $\Gamma = \Gamma_i$ with $|\Gamma_i| > 0$ and $f = f_i \in \mathcal{B}(\Gamma_i)$. By a similar argument to (3.13), and (3.23), we have that $\text{Var}[\Delta_\lambda(\Gamma_i)] = o(\lambda)$ as $\lambda \rightarrow \infty$. So, by (4.5) and the Cauchy-Schwarz inequality, to prove the lemma it suffices to show that for all λ sufficiently large

$$\text{Var} \left[\sum_{w \in \mathcal{W}_\lambda(\Gamma_i)} F_\lambda(Q_w; \Gamma_i) \right] \geq C\lambda,$$

for some $C \in (0, \infty)$.

Recall Definition 4.1. The proof now follows the idea of Avram and Bertsimas (see [2], Proposition 5). Let $M = \sum_{w \in \mathcal{W}_\lambda(\Gamma_i)} \mathbf{1}_{A_w}$. Recall the packing defined by (4.2). Index the cubes Q_w for which A_w holds by $J = \{w_1, \dots, w_M\} \subseteq \mathcal{W}_\lambda(\Gamma_i)$. Then

$$\mathbb{E}[M] = \sum_{w \in \mathcal{W}_\lambda(\Gamma_i)} \mathbb{P}(A_w) \geq \rho m_\lambda(\Gamma_i) \geq C\lambda|\Gamma_i| \geq C\lambda, \quad (4.6)$$

using (4.4) for the penultimate inequality, and the fact that $|\Gamma_i| > 0$ for the final one. As above, let \mathcal{G}_λ denote the σ -field generated by the random set $J = \{w_1, \dots, w_M\}$ and the values of $F_\lambda(Q_w; \Gamma_i)$ for $w \notin J$. Then

$$\begin{aligned} \text{Var} \left[\sum_{w \in \mathcal{W}_\lambda(\Gamma_i)} F_\lambda(Q_w; \Gamma_i) \right] &\geq \mathbb{E} \left(\text{Var} \left[\sum_{w \in J} F_\lambda(Q_w; \Gamma_i) + \sum_{w \notin J} F_\lambda(Q_w; \Gamma_i) \middle| \mathcal{G}_\lambda \right] \right) \\ &= \mathbb{E} \left(\text{Var} \left[\sum_{w \in J} F_\lambda(Q_w; \Gamma_i) \middle| \mathcal{G}_\lambda \right] \right), \end{aligned}$$

using the fact that the sum over $w \notin J$ is \mathcal{G}_λ -measurable. But by condition (ii) in Definition 4.1, the $F_\lambda(Q_w; \Gamma_i)$ for $w \in J$ are conditionally independent (under \mathcal{G}_λ), so we obtain

$$\mathbb{E} \left(\text{Var} \left[\sum_{w \in J} F_\lambda(Q_w; \Gamma_i) \middle| \mathcal{G}_\lambda \right] \right) = \mathbb{E} \sum_{w \in J} \text{Var}[F_\lambda(Q_w; \Gamma_i) | \mathcal{G}_\lambda] \geq \eta \mathbb{E}[M],$$

by condition (i) in Definition 4.1. Then by (4.6), the proof is complete. \square

Under certain extra conditions, it is the case that

$$\lambda^{-1} \text{Var}[T_i] \rightarrow \sigma_i^2, \quad (4.7)$$

for some $\sigma_i^2 \in [0, \infty)$; see [4] and [11]. Often σ_i^2 is given explicitly as an integral; however, it is often non-trivial to compute or to verify that it is strictly positive.

Under additional conditions (somewhat resembling (i) in Definition 4.1 above) it can be shown that $\sigma_i^2 > 0$. When (4.7) holds with $\sigma_i^2 > 0$ for all i , we obviously have (A1). Conditions of this type were given in [4, 13], where a form of *external* stabilization is used (which roughly speaking says that not only do Poisson points beyond the radius of stabilization for x not influence x , but also x does not influence these points). The results of [4, 13] imply that in many cases of interest (4.7) holds with $\sigma_i^2 > 0$ (given extra conditions on f_i and κ). Functionals ξ for which this holds include those associated with the total edge length of the k -nearest neighbour graph, and the total number of edges in the sphere of influence graph, plus others (see [4, 13]). Then combining (4.7) with external stabilization and the existence of moments (see Section 3 of [16] for some examples) one can obtain (2.3).

4.2 Example: the k -nearest neighbour graph

The arguments indicated above are spelled out for the particular case of the k -nearest neighbour graph in Section 3.1 of [16]. Recall that for $k \in \mathbb{N}$ and a locally finite point set $\mathcal{X} \subset \mathbb{R}^d$, the k -nearest neighbour (undirected) graph on \mathcal{X} (denoted $\text{kNG}(\mathcal{X})$) is the graph with vertex set \mathcal{X} obtained by including $\{x, y\}$ as an edge whenever $y \in \mathcal{X}$ is one of the k nearest neighbours of $x \in \mathcal{X}$, or vice versa (or both). Let $\xi(x; \mathcal{X})$ be one half the sum of the lengths in $\text{kNG}(\mathcal{X})$ incident to x . Thus (for example) we have that the total length of $\text{kNG}(\mathcal{X})$ is given by

$$\sum_{x \in \mathcal{X}} \xi(x; \mathcal{X}).$$

Suppose $\Gamma_1, \dots, \Gamma_n$ are disjoint convex or polyhedral regions. We give two examples of conditions on $\{f_i\}$ and κ which, by known results together with Theorem 2.1, yield (2.3) for this case.

First, suppose that κ is bounded away from 0 on $\cup_i \Gamma_i$. Then ξ is exponentially stabilizing and has moments of all orders. If f_i is continuous on Γ_i , then (4.7) holds with $\sigma_i^2 > 0$ (see [16], Section 3.1). Hence Theorem 2.1 applies in this case. The conditions on f_i and κ may be relaxed (see [11]), but then extra work (such as making use of the nondegeneracy argument in the present paper) is needed to show that $\sigma_i^2 > 0$.

Alternatively, suppose that κ is equal to a positive constant κ_i on each Γ_i , so that \mathcal{P}_λ is a homogeneous Poisson point process with intensity $\lambda \kappa_i > 0$ on Γ_i . Suppose that $f_i = \mathbf{1}_{\Gamma_i}$, the indicator of Γ_i , for each i . Then by the results of [13], we again have that (4.7) holds with $\sigma_i^2 > 0$, and so Theorem 2.1 holds. In this case, T_i is the total length of $\text{kNG}(\mathcal{P}_\lambda \cap \Gamma_i)$.

We conclude this section by presenting an explicit multivariate CLT of this type, derived from Theorem 2.1, for the case of the nearest-neighbour (directed) graph in one dimension. The nearest-neighbour (directed) graph on locally finite point set \mathcal{X} is the graph with vertex set \mathcal{X} obtained by including (x, y) as a (directed) edge from $x \in \mathcal{X}$ to $y \in \mathcal{X}$ when y is the nearest neighbour of x (arbitrarily breaking any ties). The required moments, regularity and stabilization conditions all follow from previous work (particularly [11, 13]), and the fact that the limiting variance is non-zero follows from an explicit calculation (which we give below) based on the general results of [11].

For a finite set $\mathcal{X} \subset (0, 1)$ and a Borel set $\Gamma \subseteq (0, 1)$, let $\mathcal{L}^\alpha(\mathcal{X}; \Gamma)$ denote the total weight of the nearest-neighbour (directed) graph on \mathcal{X} , with α -power weighted edges, counting only edges originating from points of $\mathcal{X} \cap \Gamma$. That is, if $d(x; \mathcal{X}) := d_2(x; \mathcal{X} \setminus \{x\})$ denotes the (Euclidean) distance from x to its nearest neighbour in \mathcal{X} , take

$$\xi(x; \mathcal{X}) = (d(x; \mathcal{X}))^\alpha, \tag{4.8}$$

for some fixed parameter $\alpha \in (0, \infty)$. Then

$$\mathcal{L}^\alpha(\mathcal{X}; \Gamma) = \sum_{x \in \mathcal{X} \cap \Gamma} \xi(x; \mathcal{X}).$$

For $m \in \mathbb{N}$, let $\Gamma_1, \dots, \Gamma_m$ be disjoint, finite, non-null interval subsets of \mathbb{R} . In particular, let $\pi_i = |\Gamma_i| \in (0, \infty)$ be the length of the interval Γ_i . Take $f_i = \mathbf{1}_{\Gamma_i}$. Let the underlying density κ be piecewise Borel-measurable, bounded away from 0 and from ∞ , on each interval Γ_i ; in particular, for each i set $\kappa(x) = \kappa_i(x)$ for $x \in \Gamma_i$, where $\kappa_i \in \mathcal{B}(\Gamma_i)$

and $\kappa_i(x) > 0$ for all $x \in \Gamma_i$. Consider the unmarked case (so $\mathcal{M} = \{1\}$). Then for $\lambda > 0$, \mathcal{P}_λ is a Poisson point process with intensity $\kappa_i(x)\lambda$ on each Γ_i . Using the notation of Theorem 2.1, in this set-up we have that

$$T_i = \langle \mathbf{1}_{\Gamma_i}, \mu_\lambda^\xi \rangle = \sum_{x \in \mathcal{P}_\lambda \cap \Gamma_i} \xi_\lambda(x; \mathcal{P}_\lambda) = \sum_{x \in \mathcal{P}_\lambda \cap \Gamma_i} \xi(\lambda x; \lambda \mathcal{P}_\lambda),$$

the final equality by translation-invariance. By the scaling properties (‘homogeneity’) of ξ as given by (4.8), we have

$$T_i = \sum_{x \in \mathcal{P}_\lambda \cap \Gamma_i} \xi(\lambda x; \lambda \mathcal{P}_\lambda) = \lambda^\alpha \sum_{x \in \mathcal{P}_\lambda \cap \Gamma_i} \xi(x; \mathcal{P}_\lambda) = \lambda^\alpha \mathcal{L}^\alpha(\mathcal{P}_\lambda; \Gamma_i).$$

All relevant stabilization, regularity and moments conditions are satisfied. Let \mathcal{H}_1 denote a homogeneous Poisson point process of unit intensity on $(0, 1)$, and let \mathcal{U}_n denote a binomial point process consisting of n independent uniform random points on $(0, 1)$. Then by Theorems 2.2 and 2.4 of [11], for $\alpha > 0$

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \lambda^{-1} \text{Var}[T_i] &= \lim_{\lambda \rightarrow \infty} \lambda^{2\alpha-1} \text{Var}[\mathcal{L}^\alpha(\mathcal{P}_\lambda; \Gamma_i)] \\ &= V_\alpha \int_{\Gamma_i} \kappa_i(x) dx + \left(\delta_\alpha \int_{\Gamma_i} \kappa_i(x) dx \right)^2, \end{aligned} \quad (4.9)$$

where

$$V_\alpha := \lim_{n \rightarrow \infty} n^{2\alpha-1} \text{Var}[\mathcal{L}^\alpha(\mathcal{U}_n; (0, 1))], \quad (4.10)$$

and

$$\delta_\alpha := \mathbb{E}[d(0; \mathcal{H}_1)^\alpha] + \int_{\mathbb{R}} \mathbb{E}[d(0; \mathcal{H}_1 \cup \{y\})^\alpha - d(0; \mathcal{H}_1)^\alpha] dy. \quad (4.11)$$

Let $\Gamma(\cdot)$ denote the (Euler) Gamma function, and let ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$ denote the (Gauss) hypergeometric function (see e.g. [1], Chapter 15). By (4.10) and equation (24) in [12], we have that for $\alpha > 0$

$$\begin{aligned} V_\alpha &= (4^{-\alpha} + 2 \cdot 3^{-1-2\alpha})\Gamma(1+2\alpha) - 4^{-\alpha}(3+\alpha^2)\Gamma(1+\alpha)^2 \\ &\quad + 8 \cdot \frac{6^{-\alpha-1}\Gamma(2+2\alpha)}{(1+\alpha)} {}_2F_1(-\alpha, 1+\alpha; 2+\alpha; 1/3). \end{aligned} \quad (4.12)$$

We now compute δ_α . By standard properties of the Poisson process, $D := d(0; \mathcal{H}_1)$ is distributed as an exponential random variable with parameter 2. So we have that for $\alpha > 0$

$$\mathbb{E}[D^\alpha] = \int_0^\infty 2r^\alpha \exp(-2r) dr = 2^{-\alpha} \Gamma(1+\alpha),$$

(using Euler’s Gamma integral; see e.g. 6.1.1 in [1]). By Fubini’s theorem and (4.11) we have

$$\begin{aligned} \delta_\alpha &= \mathbb{E} \left[D^\alpha - 2 \int_0^D (D^\alpha - t^\alpha) dt \right] = \mathbb{E}[D^\alpha + ((2/(1+\alpha)) - 2)D^{1+\alpha}] \\ &= 2^{-\alpha} \Gamma(1+\alpha) - \frac{2\alpha}{1+\alpha} 2^{-1-\alpha} \Gamma(2+\alpha) = 2^{-\alpha} \Gamma(1+\alpha)(1-\alpha), \end{aligned} \quad (4.13)$$

using the functional relation $\Gamma(x) = x^{-1}\Gamma(1+x)$ (see e.g. 6.1.15 in [1]) for the final equality. Of note is the fact that $\delta_1 = 0$, so that in the $\alpha = 1$ case the constant in the limiting (scaled) variance is the same in the Poisson and binomial cases. For $\alpha \neq 1$, $\delta_\alpha^2 > 0$ and the variance in the Poisson case is greater than that in the binomial case, as one expects (the Poisson process introduces additional randomness).

Also by Theorem 2.1 of [11] and equation (22) in [12], we have that for $\alpha > 0$

$$\lambda^{\alpha-1}\mathbb{E}[\mathcal{L}^\alpha(\mathcal{P}_\lambda; \Gamma_i)] \rightarrow 2^{-\alpha}\Gamma(1+\alpha) \int_{\Gamma_i} \kappa_i(x)dx,$$

as $\lambda \rightarrow \infty$. Thus we have the following application of Theorem 2.1.

Theorem 4.1 *For $m \in \mathbb{N}$, let $\Gamma_1, \dots, \Gamma_m$ be disjoint intervals in \mathbb{R} with $|\Gamma_i| = \pi_i \in (0, \infty)$. Let $\kappa(x) = \sum_{i=1}^m \kappa_i(x)\mathbf{1}_{\Gamma_i}(x)$ where, for each i , $\kappa_i \in \mathcal{B}(\Gamma_i)$ and $\kappa_i(x) > 0$ for all $x \in \Gamma_i$. Suppose $\alpha \in (0, \infty)$.*

(i) *For $1 \leq i \leq m$,*

$$\lim_{\lambda \rightarrow \infty} \lambda^{\alpha-1}\mathbb{E}[\mathcal{L}^\alpha(\mathcal{P}_\lambda; \Gamma_i)] = 2^{-\alpha}\Gamma(1+\alpha) \int_{\Gamma_i} \kappa_i(x)dx.$$

(ii) *For $1 \leq i \leq m$,*

$$\lim_{\lambda \rightarrow \infty} \lambda^{2\alpha-1}\text{Var}[\mathcal{L}^\alpha(\mathcal{P}_\lambda; \Gamma_i)] = V_\alpha \int_{\Gamma_i} \kappa_i(x)dx + \left(\delta_\alpha \int_{\Gamma_i} \kappa_i(x)dx \right)^2 =: \sigma_i^2,$$

where V_α and δ_α are given by (4.12) and (4.13) respectively.

(iii) *Given $\varepsilon > 0$, there exists $C \in (0, \infty)$ such that for all $\lambda \geq 1$,*

$$\sup_{t_1, \dots, t_m \in \mathbb{R}} \left| \mathbb{P} \left[\bigcap_{i=1}^m \left\{ \frac{\mathcal{L}^\alpha(\mathcal{P}_\lambda; \Gamma_i) - \mathbb{E}[\mathcal{L}^\alpha(\mathcal{P}_\lambda; \Gamma_i)]}{(\text{Var}[\mathcal{L}^\alpha(\mathcal{P}_\lambda; \Gamma_i)])^{1/2}} \leq t_i \right\} \right] - \prod_{i=1}^m \Phi(t_i) \right| \leq C\lambda^{\varepsilon-(1/2)}.$$

Part (iii) of Theorem 4.1 is our multivariate CLT. In the particular case of piecewise constant κ , that is $\kappa_i(x) = \kappa_i \in (0, \infty)$ for all $x \in \Gamma_i$, we have that

$$\int_{\Gamma_i} \kappa_i(x)dx = \kappa_i|\Gamma_i| = \kappa_i\pi_i,$$

and so, for example, $\sigma_i^2 = V_\alpha \kappa_i \pi_i + \delta_\alpha^2 \kappa_i^2 \pi_i^2$. Table 1 gives some values of the constants V_α , given by (4.12), and δ_α^2 , given by (4.13).

α	$1/2$	1	2	3	4
V_α	$\frac{1}{2} + \sqrt{2} \arcsin(1/\sqrt{3}) - \frac{13\pi}{32} \approx 0.094148$	$\frac{1}{6}$	$\frac{85}{108}$	$\frac{149}{18}$	$\frac{135793}{972}$
δ_α^2	$\frac{\pi}{32}$	0	$\frac{1}{4}$	$\frac{9}{4}$	$\frac{81}{4}$

Table 1: Some values of V_α and δ_α^2 .

Acknowledgements

The authors began this work while at the University of Durham, where AW was supported by an EPSRC studentship. Some of this work was carried out while AW was at the University of Bath. MP thanks the Institute for Mathematical Sciences at the National University of Singapore for its hospitality.

References

- [1] Abramowitz, M. and Stegun, I.A. (Eds.) (1965) Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series, **55**. U.S. Government Printing Office, Washington D.C.
- [2] Avram, F. and Bertsimas, D. (1993) On central limit theorems in geometrical probability, *Ann. Appl. Probab.*, **3**, 1033–1046.
- [3] Baldi, P. and Rinott, Y. (1989) On normal approximations of distributions in terms of dependency graphs, *Ann. Probab.*, **17**, 1646–1650.
- [4] Baryshnikov, Yu. and Yukich, J.E. (2005) Gaussian limits for random measures in geometric probability, *Ann. Appl. Probab.*, **15**, 213–253.
- [5] Chen, L. and Shao, Qi-Man (2004) Normal approximation under local dependence, *Ann. Probab.*, **32**, 1985–2028.
- [6] Götze, F. (1991) On the rate of convergence in the multivariate CLT, *Ann. Probab.*, **19**, 724–739.
- [7] Goldstein, L. and Rinott, Y. (1996) On multivariate normal approximations by Stein’s method, *J. Appl. Probab.*, **33**, 1–17.
- [8] Kesten, H. and Lee, S. (1996) The central limit theorem for weighted minimal spanning trees on random points, *Ann. Appl. Probab.*, **6**, 495–527.
- [9] Penrose, M. (2003) Random Geometric Graphs, *Oxford Studies in Probability*, **6**, Clarendon Press, Oxford.
- [10] Penrose, M.D. (2005) Multivariate spatial central limit theorems with applications to percolation and spatial graphs, *Ann. Probab.*, **33**, 1945–1991.
- [11] Penrose, M.D. (2005) Convergence of random measures in geometric probability. Preprint available from <http://arxiv.org/abs/math.PR/0508464>.
- [12] Penrose, M.D. and Wade, A.R. (2006) Limit theory for the random on-line nearest-neighbour graph. To appear *Random Structures Algorithms*. Preprint available from <http://arxiv.org/abs/math.PR/0603561>.
- [13] Penrose, M.D. and Yukich, J.E. (2001) Central limit theorems for some graphs in computational geometry, *Ann. Appl. Probab.*, **11**, 1005–1041.
- [14] Penrose, M.D. and Yukich, J.E. (2002) Limit theory for random sequential packing and deposition, *Ann. Appl. Probab.*, **12**, 272–301.
- [15] Penrose, M.D. and Yukich, J.E. (2003) Weak laws of large numbers in geometric probability, *Ann. Appl. Probab.*, **13**, 277–303.
- [16] Penrose, M.D. and Yukich, J.E. (2005) Normal approximation in geometric probability. In *Stein’s Method and Applications*, eds. A.D. Barbour, Louis H.Y. Chen, Lecture Notes Series, Institute for Mathematical Sciences, Vol. 5, World Scientific, Singapore.

- [17] Rinott, Y. and Rotar, V. (1996) A multivariate CLT for local dependence with $n^{-1/2} \log n$ rate and applications to multivariate graph related statistics, *J. Multivariate Anal.*, **56**, 333–350.
- [18] Stein, C. (1972) Approximate Computation of Expectations, IMS, Hayward, CA.